People-Centric Mobile Sensing with a Pragmatic Twist: from Behavioral Data Points to Active User Involvement

Jan Blom
Nokia Research Center
PSE-C, EPFL,
1015 Lausanne, Switzerland
jan.blom@nokia.com

Daniel Gatica-Perez
Idiap Research Institute
and EFPL,
Switzerland
gatica@idiap.ch

Niko Kiukkonen
Nokia Research Center
Itämerenkatu 11-13,
00180 Helsinki, Finland
niko.kiukkonen@nokia.com

ABSTRACT

Mobile phones have recently been used to collect large-scale continuous data about human behavior. This people centric sensing paradigm is useful not only from a scientific point of view: Contextual user data has pragmatic value, too. Individuals whose data is collected in such long-term people centric sensing projects can be engaged in user centric design activities aiming to generate data driven services that benefit the end user. This paper demonstrates the value of such user centric approach. In a two-stage approach, we analyse mobile phone data to extract mobile phone usage categories. We then go on to interview the participants concerning their perceptions toward context-aware services. The two stages, combined as we present here, offer a clear value in terms of providing complementary insights, both to researchers and users, about the feasibility of and the expectations about personalized mobile services.

Author Keywords
People sensing, context aware services, digital design.

ACM Classification Keywords
H5.2 User Interfaces: User Centered design.

General Terms
Data, design.

INTRODUCTION

As the penetration of mobile phones, as well as the range of sensors embedded in these devices, have reached high levels, a people centric sensing approach has emerged [1,3]. This implies the use of sensors in mobile phones and other wireless devices to collect large quantities of continuous, longitudinal data for various types of behaviors, ranging from intra- [2] to inter-personal level [3]. Computational techniques are applied to the resulting data sets, leading to the discovery of high-level attributes that emerge from the data. From a scientific point of view, this leads to new types of approaches to studying social systems, such as social networks or organizations. The capacity to collect and analyze vast amounts of socially grounded data has been recently referred to as computational social science [4].

People-centric sensing is also an industry trend. Increased sensing capabilities of mobile devices, as well as the possibility of uploading data to servers in real-time fashion, are used in context aware services. Nokia Maps [6] and Facebook Places [12] can show the present location as well as the status of members of a user's social network on a map. Hence people are now, and at once, the carriers of sensing devices and the sources and consumers of sensed events [5]. Given the fact that the sensed data features in multiple types of services and applications, a practical motivation to conduct large-scale people-centric sensing research arises, namely one related to designing digital services perceived as useful by the end users themselves.

User Centered Approach to People Sensing

In this paper, we describe the design and implementation of a people centric sensing project run by a corporate research lab [13]. Due to the corporate background, a dual motivation for the sensing activities was acknowledged from the start; both theoretical and pragmatically driven reasons provided a justification for running the research. The dual motivation led to two different perspectives to viewing the participants. In behavioral terms, sensing can lead to discovery of novel aspects of human behavior. The key value lies in the ability to capture continuous data about behavior in non-intrusive manner, consequently enabling powerful statistical techniques to be used for data analysis. This perspective was seen as theory-driven, and it could be considered as basic research contributing to current knowledge in the scientific literature.

From the user-centric perspective, the participants can be viewed as potential users of future applications and services that are based on sensing. This leads to a need to engage the participants in reflective activities, thus converging on the notion of people centric sensing, and the potential value thereof. User centered design methods [7] are applied; the participants are viewed as co-creators. They are familiarized with the nature of people sensing through the campaign. This, in turn, enables articulation of perceptions

MobileHCI 2011, Aug 30–Sept 2, 2011, Stockholm, Sweden.

and desires concerning context-aware services in an ecologically valid way.

The value of the behavioral level has been clearly demonstrated by prior research [2, 3]. This paper focuses on highlighting research falling under the user centric one. This level remains largely unexplored in prior literature concerned with people sensing, despite being important from the point of view of the evolution of mobile services. Before describing the empirical findings, the next section lays out the method used by the campaign.

CAMPAIGN METHOD

Participants and Procedure

An initial set of individuals was chosen from Lake Geneva area to serve as seeds for generating the sample. These "start nodes" were encouraged to recruit members of their social network, eventually leading to a population comprising connected individuals from a variety of backgrounds. Due to its viral nature, recruitment was not a one-off event, but a process that went on for over one year. At the time of writing this paper, the campaign population had reached a total size of 145 (66% males, 34% females). The mean age of the campaign members was 29 years, with a range between 18 and 62 years. The initial recruitment questionnaire data revealed the following occupational range: 56% employed, 9% not presently employed, 32% students, and 3% listed their status as 'other'.

We encouraged the participants to stay in the campaign for period of one year. Nokia N95 8GB smart phones were used to collect the data. Participants were instructed to use this device as their primary mobile phone.

Data Collection Technology

The data modalities enabled by the client included social interaction, location, media use, and general phone use. Social interaction was inferred from call and SMS logs and Bluetooth scanning results. Location was determined based on GPS, cellular network, and WiFi access point data. Locations visited and transportation means used could be derived from the location data. Locations where images had been taken, videos shot, or music played indicate media use. In terms of general phone use, information was received about communication (call and SMS logs), application use, and movement activity (acceleration data). No media or communication content was recorded.

The collection of the behavioral data was non-intrusive. Data uploads were automatic and occurred on a daily basis using WiFi access points. The data collection software installed on the phones provided good compromise between data amount and accuracy, and power consumption.

FINDINGS

We now illustrate how a two-stage approach was applied on the data. This section consists of two parts. We first describe the results of a factor analysis that extracts the main patterns of mobile phone usage among the participants. The data collected for this analysis stems from the non-intrusive mobile client described in the above section. In the second stage, based on the results of the factor analysis, we recruited a set of individuals for a user centric probe. The interview results are described in the second part of this section.

Use of the Phone Among Campaign Participants

The phone used by the participants allowed multiple functions to be performed, ranging from time management to multimedia content generation to Web browser and route navigation with maps. The following table summarizes the use of the key phone applications in the population for a subset of the campaign data collected between September 2009 and February 2010 (the mean number of campaign days for each participant in this subset was 105 (std=54.8)). As the joining date varied, the figures were normalized to reflect daily use. Applications tracked included: phone, SMS, camera, calendar, Web browser, and maps. For phone calls and SMS, both the number of incoming and outgoing calls/messages was factored in. For the rest of the functions, the figure is indicative of number of times the given application was visited.

	Mean	SD	Min, Max
Phone calls	2.9	2.3	.04, 11.1
SMS	3.8	6.0	.03, 54.5
Camera	0.8	1.4	0.0, 11.5
Calendar	0.16	0.3	0.0, 1.8
Browser	1.8	3.3	0.0, 25.2
Maps	0.6	1.1	0.0, 10.0

Table 1. Application use per day across the users (N=145).

Not surprisingly, the traditional communication functions of the mobile phone (voice call and SMS) were the most frequently used applications, with average daily use equaling 3 voice calls and 4 SMS. SMS use had large variation with the most active person using it 50 times per day. All 145 campaign members used both voice call and SMS features at least to some extent.

Among the investigated phone apps (camera, calendar, browser, and maps), the browser was the most popular one, with average daily use being close to 2. This reflects the recent trend of use of mobile devices for Internet access. Camera and navigation application were used less than once per day, while the calendar, perhaps surprisingly, remained the least used application, with an average use of about once a week

To obtain an understanding of differences within the population concerning application use, factor analysis with Varimax rotation was performed on these variables. Three factors emerged with eigenvalues higher than 1. In the solution, these three factors explain 72.0 % of the variance, with the following values being associated with each of the three factors: 34% (Factor 1), 20% (Factor 2) and 18% (Factor 3). Table 2 shows the component matrix across the main three factors.

	Factor		
	1	**2**	**3**
Calls	-.055	.848	.253
SMS	.068	.074	.934
Camera	.856	.104	-.132
Calendar	.345	.630	-.331
Browser	.686	.242	.072
Maps	.842	-.108	.101

Table 2. Application use per day across the users (N=145).

Factor 1 implies that some users made effective use of the multimedia related functionality of the phone, beyond voice call and SMS: Camera, maps and browser were used frequently in this category. We propose to title this category "Rich Media Users". The second factor corresponds to use patterns based on high frequency phone calls as well as calendar use. Due to the prevalence of time management, we titled this category as "Enterprise Users". The third category is associated with high frequency SMS use and lower use of phone calls. Calendar use among this segment was on the low side, relative to the rest of the population. This segment is referred to as "Texters". Further analysis of the projection of the original data on the factors shows that 41%, 27%, and 10% of the population had their strongest component along the factors Rich Media Users, Enterprise Users, and Texters, respectively.

These results show that different types of phone use patterns exist in the population, suggesting a variation in the user base of mobile phones. Further research questions originating from here relate to the self-perception of phone usage: do individuals falling under the discovered categories identify with the naming of the category? What is the "ground truth" with respect to these user types? How can we generate it or approximate it? It would also be interesting to study the impact of phone type and user interface on phone use patterns: Given a different factor and user interface, would the use of the phone differ? Are these user categories generalizable across phone types, geographic areas, cultures, etc.? In the next section, we describe how user centric design techniques were incorporated to the mining insights in order to increase our understanding of the Rich Media User segment.

Probing Rich Media Users

Four campaign members falling in the "Rich Media User" category characterized above were selected for individual interviews. The participants were relatively similar to each other: all of them were males and highly educated, ages 23-35. They had all been part of the campaign for several months by the time of recruiting them for the interviews. In line with the objectives of the study, the interviews were designed to assess perceptions toward people-centric sensing and context aware services. As the participants had first hand experience of their context being tracked on a continuous basis, the ecological validity of the findings was assumed to be on a high level. The reason for recruiting individuals from this category was because Rich Media Users were thought to be at the forefront of adopting context aware services.

The interviews consisted of two stages. The 1st stage assessed opinions toward submitting personal data to the service provider. The 2nd part familiarized the participants with the idea of context aware services, to get them to discuss the connection between people sensing and mobile services, in general. 5 service concepts were introduced to the interviewees, relating to wellness [8], traffic congestion [9], ecological lifestyle [10], location-based social networking [6] as well as pulse of city [11]. The concepts included academic and commercial examples of context aware services; they represented a broad spectrum from point of view of use motivation.

Interview findings

The interviews provided rich insights to perceptions and desires related to context aware services. Three most prevalent themes are described in further detail here. In each case, design implications are also discussed.

1: Desire for being in control of personal data

Desire to be in control of one's data emerged strongly in the interviews. Facebook was taken up as a real world example by three of four interviewees. Male 4 referred to the battle over ownership of photos of Facebook users:

"There was a big issue when Facebook decided to make pictures uploaded by users property of Facebook. People were against it and finally got right to own their own pictures."

Desire for being in control over the data collection process was observed also for the mobile scenario:

"If this is by default an app on the phone [data collection client], then maybe I will not prefer that. I should be explicitly asked whether this is fine."- Male 2

The design implications around this theme are clear: consent is needed when persuading users to take up context aware services. The philosophy of the data collection activity should be one that attributes a high sense of ownership over the data to the user herself.

2: Trade-off between privacy and personal benefits

Loss of privacy, as caused by data collection, was acknowledged by all interviewees. This was thought to be acceptable, however, if the service would bring benefits to the user, in return. For instance, Male 2 commented on Google's approach to personal data:

"Google collects my search results and customizes to my history. It's a breach of privacy but the app is so good that I cannot leave it [laughing]."

Male 4 had joined our data collection campaign because of a concrete benefit, namely opting for the same operator together with friends so as to minimize mobile expenses:

"In terms of thinking what we are loosing and what we are gaining, then privacy is what we are giving away. When we joined this [campaign], I and my friends, we chose the same operator. Then we have a lot of free talking amongst ourselves. We should have around 15 people in the campaign."

The implication here is that designers of context aware services should ask the question: how does the user gain benefit from the service? Once the answer to this question is clear, design efforts should be used to ensure crisp delivery of the message to the end users.

3: Who benefits from context aware services?

Two types of beneficiaries of people centric sensing were articulated by the interviewees: the user herself and collective entities such as municipalities or governments. The personal benefit was mostly seen in the application being able to reveal aspects of one's behavior that would be difficult to estimate otherwise. For instance, Male 4 referred to the wellness genre [8] as follows:

"I would know how much exercise I have done. I set some target, and if I see that I'm below it, it could be daily tracking of some exercise."

From the collective perspective, the wellness area led Male 1 to envision a scenario wherein government officials can access macro-level trends about exercise levels of citizens, to introduce suitable interventions when appropriate:

"We have problems with obesity. You could really get an idea how many steps people are doing and why. You could combine that with what kind of job they have. Lifestyle. I mean, what kind of transport they are using, car or train? Can you motivate them to take other types of transport?"

The third theme, related to beneficiaries, highlights an interesting future design avenue. Would mobile phone based data collection benefiting the communal bodies be an acceptable form of people centric sensing to citizens? Could an altruistic use motivation arise for context aware services, one related to contributing to the common good?

Each of the themes described above has specific design implications with regard to increasing the appeal of context aware services, at least among this user type. Needless to say, additional research is required to scale up the sample size of interviewed users. It would also be interesting to study the remaining two user categories, Enterprise Users and Texters, to find out which themes carry across the user categories, and which ones remain specific to a certain category. These are two open lines of work. Finally, if the findings can reach a more general level, a variety of design techniques could be used to generate empirically grounded concepts for context aware services.

SUMMARY AND CONCLUSIONS

This paper introduced a framework to enable both theoretically and pragmatically driven research in people-centric mobile large-scale sensing, driven by a carefully designed phone mobile data collection campaign. We presented the results of an analysis that was pragmatically driven, aiming to yield empirically grounded insights towards design of data-driven applications and services. On one hand, a factor analysis on large-scale data highlighted phone usage patterns that, while not surprising, have been so far studied mainly through qualitative techniques. Further analysis will be aimed at mining other, less evident patterns. On the other hand, our user centric probe, designed based on the automatically discovered user categories, revealed themes with clear design implications.

The empirical findings presented here are the outcome of initial research; more systematic and larger scale efforts to analyze the data and co-create with the campaign members are needed. The findings allow the conclusion, however, that computational social science methods coupled with a pragmatic mindset represent a feasible approach toward designing future context aware services.

Acknowledgements. We thank Trinh-Minh-Tri Do (Idiap) for support with data analysis.

REFERENCES

1. Campbell, A., Eisenman, S., Lane, N., Miluzzo, E., Peterson, R. People-centric urban sensing. In Proc. WICON'06, ACM Press (2006).
2. Farrahi, K., Gatica-Perez, D. What did you do today? Discovering daily routines from large scale mobile data. In Proc. MM'08, ACM Press (2008), 849-852.
3. Eagle, N., Pentland, A., and Lazer, D. Inferring social network structure using mobile phone data. *Proc. National Academy of Sciences 103*, 36 (2009), 15274-15278.
4. Lazer, D., Pentland, A., Adamic, L., Aral, S., Laszlo-Barabasi, A., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M., Jebara, T., King, G., Macy, M., Roy, D., Van Alstyne, M. Computational social science. *Science, 323*, 6 (2009), 721-723.
5. Miluzzo, E., Lane, N., Fodor, K., Peterson, R., Lu, H., Musolesi, M., Eisenman, S., Zheng, X., Campbell, A. Sensing meets mobile social networks. In Proc. SenSys'08, ACM Press (2008), 337-350.
6. Nokia Ovi Maps, http://blog.ovi.com/2010/01/22/ share-your-location-with-maps/
7. Norman, D. *The design of everyday things.* Doubleday: New York, 1998.
8. Consolvo, S., Everitt, K., Smith, I., Landay, J. Design requirements for technologies that encourage physical activity. In Proc. CHI'06, ACM Press (2006), 457-466.
9. Traffic.com, http://www.traffic.com
10. Go Green Mobile Carbon Tracker, http://www.gogreenfoundation.com/ggf_mobile.htm
11. Citysense, http://www.citysense.com/
12. Facebook Places, http://www.facebook.com/places/
13. Kiukkonen, N., Blom, J., Dousse, O., Gatica-Perez, D., and Laurila, J. Towards Rich Mobile Phone Datasets: Lausanne Data Collection Campaign. In Proc. ICPS'10, ACM Press (2010).